



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2018

Digital footprints: an emerging dimension of digital inequality

Micheli, Marina ; Lutz, Christoph ; Büchi, Moritz

Abstract: Purpose: This conceptual contribution is based on the observation that digital inequalities literature has not sufficiently considered digital footprints as an important social differentiator. The purpose of the paper is to inspire current digital inequality frameworks to include this new dimension. Design/methodology/Approach: Literature on digital inequalities is combined with research on privacy, big data and algorithms. The focus on current findings from an interdisciplinary point of view allows for a synthesis of different perspectives and conceptual development of digital footprints as a new dimension of digital inequality. Findings: Digital footprints originate from active content creation, passive participation and platform-generated data. The literature review shows how different social groups may experience systematic advantages or disadvantages based on their digital footprints. A special emphasis should be on those at the margins, for example, users of low socioeconomic background. Originality/value: By combining largely independent Research fields, the contribution opens new avenues for studying digital inequalities, including innovative methodologies to do so.

DOI: <https://doi.org/10.1108/JICES-02-2018-0014>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-167445>

Journal Article

Accepted Version

Originally published at:

Micheli, Marina; Lutz, Christoph; Büchi, Moritz (2018). Digital footprints: an emerging dimension of digital inequality. *Journal of Information, Communication and Ethics in Society*, 16(3):242-251.

DOI: <https://doi.org/10.1108/JICES-02-2018-0014>

Digital Footprints: An Emerging Dimension of Digital Inequality

Marina Micheli, Christoph Lutz, and Moritz Büchi

Abstract

Purpose – This conceptual contribution is based on the observation that digital inequalities literature has not sufficiently considered digital footprints as an important social differentiator.

Design/methodology/approach – Literature on digital inequalities is combined with research on privacy, big data, and algorithms. The focus on current findings from an interdisciplinary point-of-view allows for a synthesis of different perspectives and conceptual development of digital footprints as a new dimension of digital inequality.

Findings – Digital footprints originate from active content creation, passive participation, and platform-generated data. The literature review shows how different social groups may experience systematic advantages or disadvantages based on their digital footprints. A special emphasis should be on those at the margins, for example users of low socioeconomic background.

Originality/value – By combining largely independent research fields, the contribution opens new avenues for studying digital inequalities, including innovative methodologies to do so.

Keywords digital inequality, digital divide, digital footprints, online participation, algorithms, big data, privacy, passive participation, online platforms

Marina Micheli (m.micheli@ikmz.uzh.ch), Department of Communication and Media Research (IKMZ), University of Zurich.

Christoph Lutz (christoph.lutz@bi.no), Nordic Centre for Internet & Society, Department of Communication and Culture, BI Norwegian Business School.

Moritz Büchi (m.buechi@ikmz.uzh.ch), Department of Communication and Media Research (IKMZ), University of Zurich.

Digital footprints: an emerging dimension of digital inequality

The digital inequality literature has focused on the antecedents and effects of differentiated Internet use. More specifically, this line of research has called attention to certain online dimensions in which offline inequalities are reproduced. DiMaggio and colleagues (2004) distinguished between Internet access (technical means and autonomy of use), skill, availability of social support and variation in use, while van Dijk (2005) distinguished between motivational access, material access, skills access and usage access. Both conceptualizations show the multi-dimensionality of digital inequalities and the importance of differentiating dimensions when analyzing the phenomenon. Across all dimensions, the scientific and public policy relevance of the digital divide is based on the assumption that those who are able to effectively use the Internet might increase their social, economic, cultural, and human capital. Consequently, if users in privileged social positions have better conditions for Internet access, skills, and use, social inequalities may be exacerbated (DiMaggio et al., 2004).

In this article, we propose that predominant digital inequality frameworks should be extended to include a new dimension: digital footprints. We define digital footprints as the aggregate of data derived from the *digitally traceable behavior and online presence associated with an individual*. Not only Internet use but also individuals' digital footprints can lead to beneficial and adverse outcomes, short-term or long-term, individual or societal. What users *do* online matters; but, what *is* online about them also has consequences. As with actual footprints, digital footprints are not a complete picture of a person, but they do allow diverse inferences of varying accuracy.

We therefore argue that digital inequality scholarship should consider how digital footprints vary according to socio-demographic variables and traditional markers of inequalities. Moreover, this line of scholarship should outline potential consequences of varying digital footprints and how they might then contribute to the reproduction of social inequalities. Digital inequality research needs to explicitly account for the power imbalance between digital platforms and users in the current digital environment (Andrejevic 2014) as well as the role of big data in reproducing inequalities (Eubanks 2018; O'Neil 2016).

The increasing use of algorithms and artificial intelligence (AI) across various life domains (Beer, 2017; Gillespie, 2014) necessitates a thorough investigation of digital footprints specifically. Many algorithms rely heavily on personal data – not necessarily just online data – to automate complex tasks in the background. For instance, automated decision-making through personal data-based algorithms occurs in the context of social media content curation (Bucher, 2017), social credit scoring (Harris, 2018), recommender systems in online shopping and entertainment environments (e.g., Netflix, cf. Hallinan and Striphas, 2016), HRM and hiring practices (Mann and O’Neil, 2016), and justice systems, where algorithms attempt to predict recidivism (Dressel and Farid, 2018). Various examples, such as Microsoft’s Tay chatbot (Neff & Nagy, 2016), have demonstrated how algorithms can become problematic, raising ethical questions about transparency, accountability, bias, and discrimination (Mittelstadt et al., 2016). In particular, the issues of bias and discrimination are strongly connected to social inclusion and the representation of individuals as data subjects. Recent work has shown how systemic biases, such as racial and gender stereotypes, can be embedded into AI systems across search engines (Noble, 2018), chatbots (Schlesinger, O’Hara and Taylor, 2018), and face and voice recognition technologies widely used on social media (Howard and Borenstein, 2017). In that sense, algorithms and AI have become important topics within research on social inequalities and ethics in the digital society in general (see Cath et al., 2018).

The notion of a “digital footprint gap” was first elaborated by Robinson et al. (2015) in the context of digital inequalities over the life course, referring to data posted by adults about (unaware) children. Within this approach, the term “digital footprint gap” describes the differences in the amount of online traces between individuals or population groups. We propose to extend this concept, considering not only the quantity, but also the quality, and most importantly, the implications of online traces for reproducing inequalities. Therefore, we ask: *How do different social groups vary in their digital footprints, and subsequently, how do these quantitatively and qualitatively varying digital footprints produce outcomes that affect social inequalities?*

To a certain extent, the ability to manage one's own digital footprints successfully can be understood as a component of digital literacy. In fact, being digital media literate does not solely stand for the ability to find and critically analyze online information, but also for being able to effectively, securely and successfully use digital media in order to communicate, collaborate, share knowledge and express oneself (Iordache, Mariën, and Baelden 2017; Hargittai and Micheli 2018; van Deursen, 2010; Eshet-Alkalai, 2004). All these activities involve leaving online traces. Therefore, the ability to maximize the benefits deriving from digital footprints, while reducing the disadvantages stemming from limited, suboptimal or negative digital footprints, is certainly included into the broad notion of digital literacy. While digital footprints are not an entirely new dimension in digital inequality scholarship, we contend that research in this area has not systematically investigated this notion yet. The topic has only been partially addressed, especially by research on online privacy management and skills (e.g. Baruh, Secinti and Cemalcilar, 2017; Madden et al., 2017) and online content creation (e.g. Blank, 2013; Hargittai and Walejko, 2008; Hoffmann, Lutz and Meckel, 2015; Schradie, 2011). Research on privacy has shown that Internet skills, themselves dependent on education, strongly explain privacy protective behavior (Büchi, Just, and Latzer 2016), and thus the shaping of digital footprints. Social network sites make the task of managing privacy increasingly challenging, to the point that some users have developed a form of "apathy" and "cynicism", feeling that privacy violations are inevitable (Hargittai and Marwick 2016). Yet, understanding which personal information should not necessarily be available to others and knowing what to do about protecting such content is a type of skill that varies considerably across the population (Park, 2013). Low-income Internet users are still less likely to engage in privacy-protective strategies on social media (Madden et al., 2017) and are more likely to report having experienced problems related with their online data, such as having their reputation damaged because of something that happened online (Rainie et al. 2013).

The literature on online content creation, on the other hand, has shown that age is a decisive factor, with young users producing more content online than older users. Socio-economic status has a less clear effect: higher-educated and higher-income individuals are not necessary more likely to participate

online (Blank 2013; Micheli 2015)¹. Individual and sociodemographic factors, however, influence which specific social media platform people are more likely to participate in. The user base of each platform is structured along age, gender, ethnicity, income, and education (Hargittai 2015; Blank and Lutz 2017). Therefore, digital traces on social media are not representative for the general population. If this is not accounted for, they will generate biased findings that over-represent the experiences and opinions of a platform's prevalent sociodemographic group (Hargittai 2015; Lewis 2015; Blank 2016).

As this overview shows, research on privacy management and participation is relevant for the advancement of digital inequality research. However, it does not specifically address digital footprints. In fact, digital footprints not only correspond to data produced through active online content creation (or data that could eventually be hidden through privacy settings), but also depend on algorithmic operations and passive participation (Lutz and Hoffmann, 2017). Therefore, there is a need to incorporate studies that do not explicitly align themselves with digital inequality scholarship to fully address the relevance of this new dimension.

What are digital footprints? Moving beyond online content creation

Digital footprints are not only the product of active participation through content production and sharing but they may also be generated by algorithms and by other Internet users. Therefore, they are the sum of the data produced both by active and passive forms of participation (Casemajor et al. 2015; Lutz and Hoffmann 2017). While active participation corresponds to online content creation, “passive participation” is a new concept not fully explored within digital inequality scholarship. Our conceptualization of the term draws on Casemajor et al.'s (2015, 856) definition: “[passive participation is] engaging in a platform while being subject to processes of decision that happen outside of one's control” (p. 856). Two types of passive participation can be distinguished: data generated by platforms as a by-product of users' online behavior and data posted by other users but linked to an individual. Although discrimination is a crucial threat for

¹It is important to differentiate types of online content, such as whether it is entertainment- or civic-related (Blank 2013; C. Lutz, Hoffmann, and Meckel 2014; Schradie 2013).

both types of passively produced digital footprints, different concerns arise for the two types. For the first, surveillance and deceiving targeted advertising are particularly relevant; for the second, harm in user's reputation is more pertinent. These concerns pertain to Internet users from all socioeconomic backgrounds. However, it is still underinvestigated whether those with lower socioeconomic status are affected differently by their digital footprints.

Digital footprints as a by-product of users' online behavior

Social media platforms afford many simple user actions, such as liking, favoriting, following, or commenting, which are not seen as online content creation by most studies on Internet use, but contribute to digital footprints. Browsing histories, search queries, purchase histories and geolocation information are further types of sensitive data that, even if hidden to users, contribute to digital footprints. Platform algorithms have a pivotal function in generating such data (Sandvig et al., 2014). They not only encourage users to provide personal data, filling in profiles and forms, but also generate digital traces from every action users perform online. Notably, such data create value for social media companies thanks to data mining activities (Mayer-Schönberger and Cukier, 2013). Platforms and third-parties analyze, organize, classify and make sense of such data for behavioral predictions, surveillance, and advertising. Digital footprints, if accessible and analyzed with appropriate tools, offer an extremely accurate profile of an Internet user. For example, Kosinski, Stillwell and Graepel (2013) used Facebook Likes, an easily accessible type of digital traces, to predict personal attributes such as socio-demographics, ethnicity, sexual orientation, religious and political views, personality traits, or use of addictive substances with great accuracy.

The digital footprints produced by users' online behavior might have positive and negative repercussions for social inequalities. Since most of the data generation processes run automatically and in the background, a first issue of concern is whether Internet users are aware that platforms produce data from their "micro-acts of online participation" (Margetts et al. 2015). Research has shown, for example, that many users are not even aware that Facebook curates their News Feed and tailors posts according to their previous behavior (Rader 2014; Eslami et al. 2015; Micheli 2017). Overall, great differences exist in user

awareness and perspectives on social media data mining (Kennedy, Elgesem, and Miguel 2015). Attitudes vary along individual factors, such as age and socioeconomic status, as well as the various contexts and forms of tracking and monitoring (Kennedy, Elgesem, and Miguel 2015). For instance, studies have found that while many young people use social media actively, they are not particularly concerned about institutional privacy, that is, “how companies and third parties will use their information” (Young and Quan-Haase 2013).

Another crucial issue concerns the unequal consequences of the data collected by platforms. Algorithmically generated data can be used by platforms to propose content that is more in line with a user’s interests, but it can also enable discrimination (Edelman and Luca 2014; Rosenblat, Kneese, and boyd 2014; O’Neil 2016). Online platforms collect user data (from geo-location information to socio-demographics) to create profiles that are sold to advertisers and third-party companies. This might increase the likelihood for underprivileged users to receive poorer or even fraudulent offers (Madden et al., 2017). Users are targeted by online advertising based on their digital footprints, which has been shown to reproduce social and cultural distinctions (O’Neil 2016). In the domain of online advertising, O’Neil (2016) recounts the case of US for-profit colleges investing considerable amounts of money in online ads targeted at poor and vulnerable people. Difficult life conditions, as well as a lack of knowledge about the higher education system, make them easier to persuade to pay high tuition fees for a diploma that eventually has little value in the job market. To reach this specific population, colleges use online search advertising: platforms like Google allow advertisers to segment users according to countless attributes deduced from their search queries and online behavior (e.g., clicking on certain banners and coupons). For-profit colleges, the payday loan industry, and many other sectors use digital trace analytics for price discrimination (Valentino-DeVries, Singer-Vine, and Soltani 2012).

Digital footprints as data produced by other users

Internet users can “be participated” by other users (Casemajor et al., 2015; Lutz and Hoffmann, 2017). Examples include tagging, endorsements, ratings, and comments on the visible end of the spectrum, and searches (e.g., googling some-

one) and various data analyses on the less visible end. Such forms of passive participation may produce both desirable and profitable consequences, as well as unsolicited or annoying outcomes. Receiving ratings, likes, and shares may enhance someone's status online. The significance of data generated by other users is particularly evident for micro-celebrities, such as web influencers and YouTubers. In fact, such skilled Internet users constantly engage with their followers on several platforms with the purpose of receiving feedback. Micro-celebrities are aware that their followers' activities are fundamental for maintaining their popularity online (Khamis, Ang, and Welling 2016; Marwick 2015). The relevance of ratings and reviews is also vital for users of sharing economy platforms because both providers and consumers of sharing economy services largely rely on platforms' rating systems (Newlands, Lutz, and Fieseler 2017). Users with greater knowledge of social media platforms are better able to tailor their messages to reach the targeted audience and to maximize their visibility online. By doing so, they often leverage social media algorithms to their own advantage, receiving positive feedback from other users (Duffy, Pruchniewska, and Scolere 2017). Self-branding, intended as managing one's digital identity and improving the quality and quantity of data associated to one's profile, is a new dimension of Internet skills which has not been thoroughly investigated yet and may be associated with offline social inequalities (Hargittai and Micheli, 2018).

Data made available by other Internet users could also be undesirable and problematic. Users manage their reputation not just through privacy settings and attentive posting, but also by untagging controversial or unflattering photos, deleting posts that depict them negatively, and so on. Unwanted content posted to someone's profile is an instance of "other-generated face threats" (Litt et al., (2014, 449). Although privacy protection is often framed as an individual responsibility, both the social and technical contexts define what information is available about someone online. Users are made responsible for the behavior of people in their networks and this puts low-income users in an especially difficult condition (Madden et al., 2017; Marwick and boyd, 2014).

How to study digital footprints?

We argue that digital inequality research on digital footprints should combine different methodological approaches. Representative surveys could measure digital skills related to platform algorithms and privacy settings (Büchi et al., 2017). Qualitative interviews combined with social media profile analysis could also be a valuable method (Dubois and Ford, 2015). During interviews, respondents could discuss content they have posted, but also what has been posted by others, as well as by the platform itself. Moreover, interviews could be enriched by search engine use so that respondents could look for their digital traces and discuss the results with the interviewer (Marshall and Lindley, 2014). Interviews with social groups particularly affected by digital traces could investigate how digital footprints are perceived and enacted. Young users and micro-celebrities would be groups to scrutinize (Abidin, 2015). Such actor-focused methods could inform “social analytics”: how users “reflect upon, and adjust, their online presence and the actions that feed into it, through the use of analytics” (Couldry, Fotopoulou and Dickens, 2016, p. 119).

Beyond this, media content analyses of negative passive participation, for example in the form of doxing (Douglas 2016) and online harassment, could help case study selection. Finally, digital methods and software studies could offer useful insights to understand how platforms generate data, with implications for digital inequalities (Light, Burgess and Duguay, 2018). This also includes the critical study of algorithms (Sandvig et al., 2014) or how digital footprints influence reality construction and social order (Just and Latzer, 2016). Finally, on a macro-level, content analyses and legal assessments of platform documents, such as their terms and conditions and privacy policies, could enhance our understanding of digital footprints. For example, the analysis could focus on whether such documents contain information on the protection of certain groups, for example in terms of gender, age, or socioeconomic status.

Conclusion

In this article, we introduced the notion of digital footprints as a new dimension of digital inequality. We argued that previous digital inequality scholarship has failed to pay sufficient attention to users’ digital traces. After introducing the

initial concept of a “digital footprint gap”, as the differences in the amount of online traces between individuals or population groups (Robinson et al., 2015), we discussed the role of active content creation, algorithmically generated data as by-product of user activity, and of data posted by other users but linked to an individual (Lutz and Hoffmann, 2017). The latter two forms present interesting avenues for digital inequalities scholarship, as they challenge the notion of active Internet use and agency when it comes to digital divides. By considering algorithmic and other-created digital footprints, digital inequalities scholarship could venture into adjacent discourses and understand digital divides more holistically, theoretically, and in a more contextualized manner.

Systematic investigations of digital footprints, for example with methodologies we described in the previous section, would also allow for practical recommendations, particularly regarding inclusive platform design. The development of buildings, services, devices and websites accessible to all citizens, especially the elderly and disabled, is widely acknowledged as a fundamental prerequisite for an inclusive society (Clarkson et al., 2013). In the same vein, online platforms should be designed not only to be accessible, but also to prevent the occurrence of a “digital footprint gap”. In particular, under-privileged and under-represented groups could be given more voice, while those groups that are particularly vulnerable to harassment or exploitation through their digital footprints could be better protected. Simple and transparent alert mechanisms on social media are an example of design implementation in this regard, as well as specific privacy enhancing technologies (D’Acquisto et al. 2015) defined within the “privacy by design” socio-technical approach (Cavoukian, 2009; Schaar, 2010).

Digital footprints as an emerging dimension of digital inequality connect to the broader societal and academic discourse on the mutual dependencies of society and technology. While digital inequality scholarship has addressed this nexus by analyzing the mechanisms between life chances and the purposeful use of ICTs, we have argued for the inclusion of digital footprints in the analysis of what ultimately concerns informational and social justice.

References

- Abidin, C. (2015), "Communicative intimacies: influencers and perceived interconnectedness", *Ada: A Journal of Gender, New Media, and Technology*, Vol. 8, available at <http://adanewmedia.org/2015/11/issue8-abidin/>
- Andrejevic, M. (2014), "The big data divide", *International Journal of Communication*, Vol. 8, pp. 1673-1689.
- Baruh, L., Secinti, E. and Cemalcilar, Z. (2017), "Online privacy concerns and privacy management: a meta-analytical review", *Journal of Communication*, Vol. 67 No. 1, pp. 26-53.
- Beer, D. (2017), "The social power of algorithms", *Information, Communication & Society*, Vol. 20 No. 1, pp. 1-13.
- Blank, G. (2013), "Who creates content? Stratification and content creation on the Internet", *Information, Communication & Society*, Vol. 16 No. 4, pp. 590-612.
- Blank, G., & Lutz, C. (2017), "Representativeness of Social Media in Great Britain: Investigating Facebook, LinkedIn, Twitter, Pinterest, Google+, and Instagram", *American Behavioral Scientist*, Vol. 61 No 7, pp. 741-756.
- Bucher, T. (2017), "The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms", *Information, Communication & Society*, Vol. 20 No 1, pp. 30-44.
- Büchi, M., Just, N. and Latzer, M. (2017), "Caring is not enough: the importance of Internet skills for online privacy protection", *Information, Communication & Society*, Vol. 20 No. 8, pp. 1261-1278.
- Casemajor, N., Couture, S., Delfin, M., Goerzen, M. and Delfanti, A. (2015), "Non-participation in digital media: toward a framework of mediated political action", *Media, Culture & Society*, Vol. 37 No. 6, pp. 850-866.
- Cath, C., Zimmer, M., Lomborg, S., and Zevenbergen, B. (2018). Association of Internet Researchers (AoIR) Roundtable Summary: Artificial Intelligence and the Good Society Workshop Proceedings. *Philosophy & Technology*, Vol. 31 No. 1, pp. 155-162.
- Cavoukian, A. (2009), "Privacy by Design – Take the Challenge", Information and Privacy Commissioner, Toronto, Ontario.
- Clarkson, P. J., Coleman R., Keates S., and Lebbon C. (2013), *Inclusive Design: Design for the Whole Population*, Springer Science & Business Media.
- Couldry, N., Fotopoulou, A. and Dickens, L. (2016), "Real social analytics: a contribution towards a phenomenology of a digital world", *The British Journal of Sociology*, Vol. 67 No. 1, pp. 118-137.
- D'Acquisto, G., Domingo-Ferrer J., Kikiras P., Torra V., de Montjoye Y., and Bourka A. (2015), *Privacy by Design in Big Data: An Overview of Privacy Enhancing Technologies in the Era of Big Data Analytics*, ENISA: European Union Agency for Network and Information Security, <https://doi.org/10.2824/641480>.
- Deursen, Alexander J.A.M. van. (2010), "Internet Skills: Vital Assets in an Information Society", Enschede, the Netherlands: University of Twente.
- DiMaggio, P., Hargittai, E., Celeste, C. and Shafer, S. (2004), "From unequal access to differentiated use: a literature review and agenda for research on digital inequality", in Necker-man, K. (Ed.), *Social inequality*, Russell Sage Foundation, New York, NY, pp. 355-400.
- Douglas, D. M. (2016), "Doxing: a conceptual analysis", *Ethics and Information Technology*, Vol. 18 No. 3, pp. 199-210.

- Dressel, J. and Farid, H. (2018), "The accuracy, fairness, and limits of predicting recidivism", *Science Advances*, Vol. 4 No. 1, eaao5580, pp. 1-5.
- Dubois, E. and Ford, H. (2015), "Trace interviews: an actor-centered approach", *International Journal of Communication*, Vol. 9, pp. 2067-2091.
- Duffy, B. E., Pruchniewska, U. and Scolere, L. (2017), "Platform-specific self-branding: imagined affordances of the social media ecology", in *Proceedings of the 8th International Conference on Social Media & Society*, ACM, New York, NY, pp. 1-9.
- Edelman, B. G. and Luca, M. (2014), "Digital discrimination: the case of Airbnb.com", *SSRN Electronic Journal*, available at <https://papers.ssrn.com/abstract=2377353>
- Eshet-Alkalai, Y. (2004), "Digital Literacy: a Conceptual Framework for Survival Skills in the Digital Era", *Journal of Educational Multimedia and Hypermedia*, Vol. 13 No 1, pp. 93-106.
- Eslami, M., Rickman, A., Vaccaro, K., Aleyasen, A., Vuong, A., Karahalios, K., ... Sandvig, C. (2015), "I always assumed that I wasn't really that close to [her]': reasoning about invisible algorithms in news feeds", in *CHI'15: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ACM, New York, NY, pp. 153-162.
- Eubanks, V. (2018), *Automating inequality: How high-tech tools profile, police, and punish the poor*, St. Martin's Press, New York, NY.
- Gillespie, T. (2014), "The relevance of algorithms", in Gillespie T., Boczkowski, P.J. and Foot, K.A. (Eds.), *Media Technologies: Essays on Communication, Materiality, and Society*, MIT Press, Cambridge, MA, pp. 167-194.
- Hallinan, B. and Striphas, T. (2016), "Recommended for you: The Netflix Prize and the production of algorithmic culture", *New Media & Society*, Vol. 18 No. 1, pp. 117-137.
- Hargittai E. and Micheli M. (2018), "Internet skills and why they matter", In Graham M. and Dutton W.H. (Eds.) *Society and the Internet. How Networks of Information and Communication Are Changing Our Lives. Second Edition*. Oxford University Press, Oxford, UK.
- Hargittai, E. (2015), "Is Bigger Always Better? Potential Biases of Big Data Derived from Social Network Sites", *The ANNALS of the American Academy of Political and Social Science*, Vol. 659 No 1, pp. 63-76.
- Hargittai, E. and Marwick, A.E. (2016), "What can I really do?' Explaining the privacy paradox with online apathy", *International Journal of Communication*, Vol. 10, pp. 3737-3757.
- Hargittai, E. and Walejko, G. (2008), "The participation divide: content creation and sharing in the digital age", *Information, Communication & Society*, Vol. 11 No. 2, pp. 239-256.
- Harris, J. (2018), "The tyranny of algorithms is part of our lives: soon they could rate everything we do", *The Guardian*, 5 March, available at <https://www.theguardian.com/commentis-free/2018/mar/05/algorithms-rate-credit-scores-finances-data>
- Hoffmann, C.P., Lutz, C. and Meckel, M. (2015), "Content creation on the Internet: a social cognitive perspective on the participation divide", *Information, Communication & Society*, Vol. 18 No. 6, pp. 696-716.
- Hoffmann, C.P., Lutz, C. and Ranzini, G. (2016), "Privacy cynicism: a new approach to the privacy paradox", *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, Vol. 10 No. 4, available at <https://cyberpsychology.eu/article/view/6280/5888>
- Howard, A. and Borenstein, J. (2017), "The ugly truth about ourselves and our robot creations: The problem of bias and social inequity", *Science and Engineering Ethics*, Vol. advance online publication, available at <https://doi.org/10.1007/s11948-017-9975-2>

- Iordache, C., Mariën, I., & Baelden, D. (2017), "Developing Digital Skills and Competences: A Quick-Scan Analysis of 13 Digital Literacy Models", *Italian Journal of Sociology of Education*, Vol. 9 No. 1, pp. 6–30.
- Just, N. and Latzer, M. (2016), "Governance by algorithms: reality construction by algorithmic selection on the Internet", *Media, Culture & Society*, Vol. 39 No. 2, pp. 238–258.
- Kennedy, H., Elgesem, D., & Miguel, C. (2017), "On fairness: User perspectives on social media data mining", *Convergence*, Vol. 23, No. 3, pp. 270–288.
- Khamis, S., Ang, L. and Welling, R. (2016), "Self-branding, 'micro-celebrity' and the rise of social media influencers", *Celebrity Studies*, Vol. 8 No. 2, pp. 191–208.
- Kosinski, M., Stillwell, D. and Graepel, T. (2013), "Private Traits and attributes are predictable from digital records of human behavior", *Proceedings of the National Academy of Sciences*, Vol. 110 No 15, pp. 5802–5805.
- Lewis, K. (2015), "Three fallacies of digital footprints", *Big Data & Society*, Vol. 2, pp. 1–4.
- Light, B., Burgess, J. and Duguay, S. (2018), "The walkthrough method: an approach to the study of apps", *New Media & Society*. Vol. 20 No. 3, pp. 881–900.
- Litt, E., Spottswood, E., Birnholtz, J., Hancock, J. T., Smith, M. E. and Reynolds, L. (2014), "Awkward encounters of an 'other' kind: collective self-presentation and face threat on Facebook", in *CSCW'14: Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, ACM, New York, NY, pp. 449–460.
- Lutz, C. and Hoffmann, C. P. (2017), "The dark side of online participation: exploring non-, passive and negative participation", *Information, Communication & Society*, Vol. 20 No. 6, pp. 876–897.
- Lutz, C., Hoffmann, C. P. and Meckel, M. (2014), "Beyond just politics: a systematic literature review of online participation", *First Monday*, Vol. 19 No. 7, available at <http://firstmonday.org/article/view/5260/4094>
- Madden, M., Gilman, M. E., Levy, K. E. and Marwick, A. E. (2017), "Privacy, poverty and big data: a matrix of vulnerabilities for poor Americans", *SSRN Electronic Journal*, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2930247
- Mann, G. and O'Neil, C. (2016), "Hiring algorithms are not neutral", *Harvard Business Review*, 9 December, available at <https://hbr.org/2016/12/hiring-algorithms-are-not-neutral>
- Margetts, H., John, P., Hale, S. and Yasserli, T. (2015), *Political turbulence: how social media shape collective action*, Princeton University Press, Princeton, NJ.
- Marshall, C. C. and Lindley, S. E. (2014), "Searching for myself: motivations and strategies for self-search", in *CHI'14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, New York, NY, pp. 675–684.
- Marwick, A. E. (2015), *Status update: celebrity, publicity, and branding in the social media age*, Yale University Press, New Haven, CT.
- Marwick, A. E. and boyd, d. (2014), "Networked privacy: how teenagers negotiate context in social media", *New Media & Society*, Vol. 16 No. 7, pp. 1051–1067.
- Marwick, A. E., Fontaine, C., & boyd, d. (2017), "'Nobody sees it, nobody gets mad': social media, privacy, and personal responsibility among low-SES youth". *Social Media + Society*, Vol. 3 No. 2, available at <http://journals.sagepub.com/doi/abs/10.1177/2056305117710455>
- Mayer-Schönberger, V. and Cukier, K. (2013), *Big data: A revolution that will transform how we live, work, and think*, Houghton Mifflin Harcourt, New York, NY.

- Micheli, M. (2015), "What is new in the digital divide? Understanding Internet use by teenagers from different social backgrounds", in *Communication and Information Technologies Annual*, Vol. 10, Emerald, Bingley, UK, pp. 55-87.
- Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S. and Floridi, L. (2016), "The ethics of algorithms: Mapping the debate", *Big Data & Society*, Vol. 3 No. 2, pp. 1-21.
- Neff, G. and Nagy, P. (2016), "Talking to bots: Symbiotic agency and the case of Tay", *International Journal of Communication*, Vol. 10, pp. 4915-4931.
- Newlands, G., Lutz, C. and Fieseler, C. (2018), "European perspectives on power in the sharing economy", *SSRN Electronic Journal*, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3046473
- Noble, S.U. (2018), *Algorithms of oppression: How search engines reinforce racism*. NYU Press, New York, NY.
- O'Neil, C. (2016), *Weapons of math destruction: how big data increases inequality and threatens democracy*, Crown/Archetype, New York, NY.
- Rader, E. J. (2014), "Awareness of behavioral tracking and information privacy concern in Facebook and Google", in *SOUPS*, pp. 51-67, available at https://www.usenix.org/sites/default/files/soups14_proceedings.pdf
- Rainie, L., Kiesler, S., Kang, R., & Madden, M. (2013), "Anonymity, privacy, and security online", PEW available at <http://www.pewinternet.org/2013/09/05/anonymity-privacy-and-security-online/>
- Robinson, L., Cotten, S. R., Ono, H., Quan-Haase, A., Mesch, G., Chen, W., Schulz, J., Hale, T. J. and Stern, M. J. (2015), "Digital inequalities and why they matter", *Information, Communication & Society*, Vol. 18 No. 5, pp. 569-582.
- Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2014), "Auditing algorithms: research methods for detecting discrimination on internet platforms", paper presented at the *Data and discrimination: converting critical concerns into productive inquiry* preconference of the 64th Annual Meeting of the International Communication Association, May 22, 2014, Seattle, WA.
- Schaar, P. (2010), "Privacy by Design", *Identity in the Information Society*, Vol. 3 No 2, pp. 267-74.
- Schlesinger, A., O'Hara, K.P. and Taylor, A.S. (2018), "Let's talk about race: Identity, chatbots, and AI", in *CHI'18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ACM, New York, NY, pp. 1-14.
- Schradie, J. (2011), "The digital production gap: the digital divide and Web 2.0 collide", *Poetics*, Vol. 39 No. 2, pp. 145-168.
- Schradie, J. (2013), "The digital production gap in Great Britain: how sampling, mechanisms, and theory matter with digital inequality", *Information, Communication & Society*, Vol. 16 No. 6, pp. 989-998.
- Valentino-DeVries, J., Singer-Vine, J., & Soltani, A. (2012), "Websites vary prices, deals based on users' information", *Wall Street Journal*, available at <https://www.wsj.com/articles/SB10001424127887323777204578189391813881534>
- Van Dijk, J. (2005), *The deepening divide: inequality in the Information Society*, Sage, London.
- Young, A. L., & Quan-Haase, A. (2013), "Privacy Protection Strategies on Facebook: The Internet Privacy Paradox Revisited", *Information, Communication & Society*, Vol. 16, No. 4, pp. 479-500.